The Rényi Divergence and Security Proofs

Thomas Prest
Introduction

1 Introduction

Theory

2 Theory
   ① The Rényi Divergence
   ② Three useful lemmas
   ③ Framework for proving stuff

Practice

3 Practice
   ① Application 1: Security of a Sampler from [MW17]
   ② Application 2: Revisiting the Table Approach
   ③ Application 4: Standard Deviation of Trapdoor Samplers
   ④ Application 5: Precision of Trapdoor Samplers

Conclusion

4 Conclusion
   ① Quick Summary
   ② Open Questions
What is the Rényi divergence and why use it?

Security proofs involving distributions:

- **The standard approach:** use the statistical distance $\Delta$.
  - Take a hard problem relying on some ideal distribution $Q$,
  - Replace $Q$ by a “real-life” distribution $P$,
  - If $\Delta(P, Q)$ is small enough, we win: the problem is still hard.

- **Lattice-based cryptography:** often relevant to replace SD by Rényi divergence.
  - Sharper parameters [LSS14, LPSS14, BLL+15, BGM+16, Pre17, HLS17]
  - KEMs distributions [ADPS16, BCD+16]
  - Reduction between LWE problems [AD17]

This presentation:

1. Formalize and optimize the use of the Rényi divergence in security proofs ⇒ Section 2.
2. More applications of the Rényi divergence to lattice-based cryptography ⇒ Section 3.
3. A brief discussion on open problems ⇒ Section 4.

Based on [Pre17].
Theory

1 Introduction

2 Theory
   ① The Rényi Divergence
   ② Three useful lemmas
   ③ Framework for proving stuff

3 Practice

4 Conclusion
The Rényi Divergence

**Definition.** For $a \in (1, +\infty)$, the Rényi divergence between two distributions $\mathcal{P}, \mathcal{Q}$ is

$$R_a(\mathcal{P} \parallel \mathcal{Q}) = \left( \sum_{x \in \text{Supp}(\mathcal{P})} \frac{\mathcal{P}(x)^a}{\mathcal{Q}(x)^{a-1}} \right)^{\frac{1}{a-1}}$$

**Motivation.** We consider a cryptographic scheme doing $q$ queries to a distribution $\mathcal{D}_i$ ($i \in \{0, 1\}$), we note $\varepsilon_i$ the probability of an event breaking the scheme.

- With the statistical distance:

  $$\varepsilon_0 \geq \varepsilon_1 - q\Delta(\mathcal{D}_1, \mathcal{D}_0)$$

  $\Delta \leq 2^{-\lambda} \implies \text{we win}$

- With the Rényi divergence:

  $$\varepsilon_0 \geq \varepsilon_1^{\frac{a}{a-1}} / R_a(\mathcal{D}_1 \parallel \mathcal{D}_0)^q$$

  $(a \geq \lambda) \& (\log R_a \leq 1/q) \implies \text{we win}$

**Observation.** For “equal” values ($\log R_a \approx \Delta$), Rényi divergence is more interesting when $q \ll 2^\lambda$ [BLL⁺15]. And typically:

- $128 \leq \lambda \leq 256$
- $1 \leq q \leq 2^{64}$
### The first and second lemmas

1. **Tailcut.** Let $\delta > 0$ such that $\frac{D_\delta}{D} \leq 1 + \delta$. For $a \in (1, \infty]$:

$$R_a(D_\delta || D) \leq (1 + \delta)^{a/a-1}$$

*Example:* $D_\delta$ is a tailcut of $D$ (discard a set $S$ such that $D(S) \leq \delta$).

2. **Relative error.** Suppose $\text{Supp}(D_\delta) = \text{Supp}(D)$. Let $\delta > 0$ such that $1 - \delta \leq \frac{D_\delta}{D} \leq 1 + \delta$. For $a \in (1, \infty)$:

$$R_a(D_\delta || D) \leq \left(1 + \frac{a(a-1)\delta^2}{2(1-\delta)^{a+1}}\right)^{\frac{1}{a-1}} \sim 1 + \frac{a\delta^2}{2}$$

*Example:* $D_\delta$ implements $D$ with finite precision (relative error $\delta$).
The third lemma

The max-log distance. Introduced in [MW17].

For two distributions $P$ and $Q$ over the same support $S$:

$$\Delta_{ML}(P, Q) = \max_{x \in S} |\log P(x) - \log Q(x)|$$

Unlike the Rényi divergence, it is a distance, so it verifies the:

- Triangle inequality: $\Delta_{ML}(P, R) \leq \Delta_{ML}(P, Q) + \Delta_{ML}(Q, R)$
- Symmetry: $\Delta_{ML}(P, Q) = \Delta_{ML}(Q, P)$

[MW17] essentially states that $\Delta_{ML} \leq 2^{-\lambda/2} \Rightarrow$ we win.

---

1Actually similar to the differential privacy.
The third lemma

**The max-log distance.** Introduced in [MW17].\(^1\)

For two distributions \(\mathcal{P}\) and \(\mathcal{Q}\) over the same support \(S\):

\[
\Delta_{ML}(\mathcal{P}, \mathcal{Q}) = \max_{x \in S} |\log \mathcal{P}(x) - \log \mathcal{Q}(x)|
\]

Unlike the Rényi divergence, it is a distance, so it verifies the:

- Triangle inequality: \(\Delta_{ML}(\mathcal{P}, \mathcal{R}) \leq \Delta_{ML}(\mathcal{P}, \mathcal{Q}) + \Delta_{ML}(\mathcal{Q}, \mathcal{R})\)
- Symmetry: \(\Delta_{ML}(\mathcal{P}, \mathcal{Q}) = \Delta_{ML}(\mathcal{Q}, \mathcal{P})\)

[MW17] essentially states that \(\Delta_{ML} \leq 2^{-\lambda/2} \Rightarrow\) we win.

3. **A reverse Pinsker inequality.** For two distributions \(\mathcal{P}, \mathcal{Q}\) of common support, we have:

\[
R_\alpha(\mathcal{P}||\mathcal{Q}) \leq \left( 1 + \frac{a(a-1)(e^{\Delta_{ML}(\mathcal{P}, \mathcal{Q})} - 1)^2}{2(2 - e^{\Delta_{ML}(\mathcal{P}, \mathcal{Q})})^{a+1}} \right)^{\frac{1}{a-1}} \sim_{\Delta_{ML} \to 0} 1 + \frac{a\Delta_{ML}(\mathcal{P}, \mathcal{Q})^2}{2}
\]

Consequence: Instead of \(\Delta_{ML} \leq 2^{-\lambda/2}\), we only need \(\Delta_{ML} \leq \frac{1}{\sqrt{\lambda q}}\).

\(^1\)Actually similar to the differential privacy.
Framework for using the Rényi Divergence

1. Take your favourite scheme
2. Set more aggressive parameters:
   1. First, try to apply the relative error lemma (the most powerful)
   2. Wherever it doesn’t work, apply either the tailcut lemma or the reverse Pinsker’s inequality
   ⚠ Taking $R_a \leq 1 + \frac{1}{q}$ is sufficient.
   ⚠ Taking $\alpha \geq \lambda$ gives tight, efficient proofs.
3. Goto step 1

⚠ These arguments are only valid for search problems!
For decision problems, achieving the same efficiency is still open.

⚠ In the rest of this presentation, we assume $q \leq 2^{64}$. 
Practice

1. Introduction

2. Theory

3. Practice
   1. Application 1: Security of a Sampler from [MW17]
   2. Application 2: Revisiting the Table Approach
   3. Application 4: Standard Deviation of Trapdoor Samplers
   4. Application 5: Precision of Trapdoor Samplers

4. Conclusion
Application 1: Security of a Sampler from [MW17]

**Context.** A new sampler over $\mathbb{Z}$ was introduced in [MW17].

**Previous works.** [MW17] perform a max-log distance-based analysis of the sampler. They find that

$$64 \text{ bits of precision} \Rightarrow \Delta_{\text{ML}} \leq 2^{-50} \Rightarrow \text{About 100 bits of security}$$

**This work.** We use the reverse Pinsker’s inequality:

$$64 \text{ bits of precision} \Rightarrow \Delta_{\text{ML}} \leq 2^{-50} \Rightarrow R_a \leq 1 + 2^{-96} \Rightarrow 256 \text{ bits of security, even with up to } 2^{94} \text{ queries}$$

We gain this much security for free.
No knowledge about the sampler is required.
Application 2: Revisiting the Table Approach

**Context.** We study the use of precomputed tables for sampling discrete distributions – typically, (pseudo)Gaussians.

**Previous works.** Existing approaches [Pei10, PDG14, DG14] require high precision ($\geq \lambda/2$) and/or floating-point arithmetic.

**This work.** We propose a simple approach which requires less than 64 bits of fixed precision in practice.
The classical CDF-table approach

Let $\mathcal{D}$ be a distribution over $\mathbb{N}$ that we want to sample from. We suppose we have a precomputed table of CDF$_{\mathcal{D}}$ defined over $\mathbb{N}$ by:

$$\text{CDF}_\mathcal{D}(z) = \sum_{i\leq z} \mathcal{D}(i)$$

**Algorithm 1** CDF sampler

**Require:** A precomputed table of CDF$_\mathcal{D}$

1: $z \leftarrow 0$
2: $u \leftarrow [0, 1]$ uniformly
3: **while** $u \geq \text{CDF}_\mathcal{D}(z)$ **do**
4: $z \leftarrow z + 1$
5: **Return** $z$

Suppose we want to sample a half-Gaussian $D^+_\sigma$.

- **Statistical distance-based analysis.** We need to store about:
  - $\sigma \cdot \sqrt{2\lambda}$ values,
  - With a precision $\lambda$.
- **Rényi Divergence-based analysis.** We need to store about:
  - $\sigma \cdot \sqrt{2q}$ values,
  - With a precision $\lambda$. But we prefer/expect $\log_2(q)$ or $\log_2(q)/2$!
The CoDF sampler

**Our solution.** We use a “Rényi divergence-friendly“ table. This requires a different algorithm. We define the conditional density function of $\mathcal{D}$ by:

$$\text{CoDF}_{\mathcal{D}}(z) = \mathcal{D}(z) / \sum_{i \geq z} \mathcal{D}(i)$$

---

**Algorithm 2 CoDF sampler**

**Require:** A precomputed table of $\text{CoDF}_{\mathcal{D}}$

**Ensure:** $z \leftarrow \mathcal{D}$

1. $z \leftarrow 0$
2. $u \leftarrow [0, 1]$ uniformly
3. **while** $u \geq \text{CoDF}_{\mathcal{D}}(z)$ **do**
   - $z \leftarrow z + 1$
   - $u \leftarrow [0, 1]$ uniformly
4. **Return** $z$

Suppose we want to sample a half-Gaussian $D^+_\sigma$.

», **Rényi Divergence-based analysis.** We need to store about:

- $\sigma \cdot \sqrt{2q}$ values,
- With a precision $\log_2(q)/2!$
Example and Conclusion

**Gain in theory:**

- CDF+SD approach: $\sigma \cdot \sqrt{2\lambda}$ values with precision $\lambda$
- CoDF+RD approach: $\sigma \cdot \sqrt{2q}$ values with precision $\log_2(q)/2$

**A practical example:** the distribution $D^+_{\mathbb{Z},0.85\ldots}$ from [DDLL13].

- CDF+SD approach: 20 elements of 266 bits each $\Rightarrow \approx 5300$ bits.
- CoDF+RD approach: 11 elements of 53 bits each $\Rightarrow \approx 600$ bits.

**Conclusion:**

- Both in theory and practice, we gain an order of magnitude.
- Requires only standard (64 bits) fixed-point arithmetic.
- Highly composable with other table-based techniques.
Application 4: Standard Deviation of Trapdoor Samplers

**Context.** Trapdoor sampling allows to sample a discrete Gaussian $D_{\Lambda(B),\sigma,c}$.  

- Allows hash-and-sign, IBE [GPV08], standard model signatures [CHKP10, Boy10], hierarchical IBE [CHKP10, ABB10a, ABB10b], attribute-based encryption [Boy13, BGG+14] and so on.  
- Current algorithms [Kle00, GPV08, Pei10, MP12, DP16] heavily rely on floating-point arithmetic.

**This work.** Two axes of improvement for trapdoor samplers:  

1. Squeezing the standard deviation  
2. Reducing the required precision  

These had critical impacts for the signature scheme Falcon [PFH+17].
Our test subject: Klein's sampler

**Algorithm 3** Klein\(_{L,\sigma}(t)\)

**Require:** \(\sigma \geq \eta \epsilon(\mathbb{Z}^n) \cdot \|B\|_{GS}\), the GSO \(B = L \cdot \tilde{B}\), values \(\sigma_j = \sigma / \|\tilde{b}_j\|\), a target \(t\)

**Ensure:** A vector \(z\) such that \(zB \leftarrow D_{\Lambda(B),\sigma,tB}\)

\[
\begin{align*}
\text{for } j = n, \ldots, 1 & \text{ do} \\
& c_j \leftarrow t_j + \sum_{i>j} (t_j - z_j)L_{ij} \\
& z_j \leftarrow D_{z,\sigma_j,c_j} \\
\end{align*}
\]

**return** \(z\)

1. \(\sigma\) too large \(\Rightarrow\) Klein\(_{L,\sigma}\) is useless in a cryptographic context.
2. \(\sigma\) too small \(\Rightarrow\) Klein\(_{L,\sigma}\) does not behave like a perfect Gaussian.

So \(\sigma\) must be small but the output of Klein\(_{L,\sigma}\) must still look like a Gaussian.
The adequate value for $\sigma$ is at the intersection of the hardness curve (constraint 1) and the SD/KLD/RD curve (constraint 2).

$\Rightarrow$ A Rényi divergence-based analysis proves to be much more efficient than an SD/KLD-based one.

$\Rightarrow$ Interesting fact: in practice, $\sigma$ is not conditioned by $\lambda$ but by $q$.

In practice, we gain about 30 bits of security (compared to the SD).
Application 5: What about the precision?

- **With the SD**: $\lambda$ bits of precision
- **With the KLD [LP15]**: $\lambda/2$ bits of precision
- **With the RD [Pre17]**: $\log_2 q/2$ bits of precision
Conclusion

1 Introduction
2 Theory
3 Practice
4 Conclusion
   ① Quick Summary
   ② Open Questions
The current state of affairs.

- Rényi divergence is a powerful tool, but not easy to use.
- With the reverse Pinsker’s inequality, the fact that the Rényi divergence is not a distance is no longer a problem.
- We can have much better parameters if these conditions are met:
  - Limited number of queries
  - Search problems
  - A bit of luck
- These results are generic (not limited to lattice-based cryptography).
Interesting questions IMHO.

- When is the Rényi divergence worse than the statistical distance?
- Applications outside lattice-based cryptography?
- Application to theoretical LBC rather than “production line” LBC?
- Achieve a similar efficiency for decision problems?
Interesting questions IMHO.

» When is the Rényi divergence worse than the statistical distance?
» Applications outside lattice-based cryptography?
» Application to theoretical LBC rather than “production line” LBC?
» Achieve a similar efficiency for decision problems?

Thanks!
Shweta Agrawal, Dan Boneh, and Xavier Boyen. Efficient lattice (H)IBE in the standard model.
In Gilbert [Gil10], pages 553–572.

Shweta Agrawal, Dan Boneh, and Xavier Boyen. Lattice basis delegation in fixed dimension and shorter-ciphertext hierarchical IBE.
In Rabin [Rab10], pages 98–115.

Martin R. Albrecht and Amit Deo. Large modulus ring-LWE ≥ module-LWE.
In Takagi and Peyrin [TP17], pages 267–296.


Joppe W. Bos, Craig Costello, Léo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take off the ring! Practical, quantum-secure key exchange from LWE.

Dan Boneh, Craig Gentry, Sergey Gorbunov, Shai Halevi, Valeria Nikolaenko, Gil Segev, Vinod Vaikuntanathan, and Dhinakaran Vinayagamurthy.
Fully key-homomorphic encryption, arithmetic circuit ABE and compact garbled circuits.
In Nguyen and Oswald [NO14], pages 533–556.

Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen.
On the hardness of learning with rounding over small modulus.

Shi Bai, Adeline Langlois, Tancrede Lepoint, Damien Stehlé, and Ron Steinfeld.
Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance.
Xavier Boyen.
Lattice mixing and vanishing trapdoors: A framework for fully secure short signatures and more.

Xavier Boyen.
Attribute-based functional encryption on lattices.

David Cash, Dennis Hofheinz, Eike Kiltz, and Chris Peikert.
Bonsai trees, or how to delegate a lattice basis.
In Gilbert [Gil10], pages 523–552.

Léo Ducas, Alain Durmus, Tancrède Lepoint, and Vadim Lyubashevsky.
Lattice signatures and bimodal Gaussians.


Léo Ducas and Thomas Prest. Fast fourier orthogonalization.
Henri Gilbert, editor. 

Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. 
Trapdoors for hard lattices and new cryptographic constructions. 

Andreas Hülsing, Tanja Lange, and Kit Smeets. 
Rounded gaussians – fast and secure constant-time sampling for lattice-based crypto. 

Philip N. Klein. 
Finding the closest lattice vector when it’s unusually close. 

Vadim Lyubashevsky and Thomas Prest.


Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In Nguyen and Oswald [NO14], pages 239–256.


Daniele Micciancio and Michael Walter. Gaussian sampling over the integers: Efficient, generic, constant-time.


Thomas Prest.
Sharper bounds in lattice-based cryptography using the Rényi divergence.
In Takagi and Peyrin [TP17], pages 347–374.

Tal Rabin, editor.

Tsuyoshi Takagi and Thomas Peyrin, editors.